FAVARD SEPARATION METHOD FOR ALMOST PERIODIC STOCHASTIC DIFFERENTIAL EQUATIONS

ZHENXIN LIU AND WENHE WANG

ABSTRACT. Favard separation method is an important means to study almost periodic solutions to linear differential equations; later, Amerio applied Favard's idea to nonlinear differential equations. In this paper, by appropriate choosing separation and almost periodicity in distribution sense, we obtain the Favard and Amerio type theorems for stochastic differential equations.

1. INTRODUCTION

The theory of almost periodic functions was founded by Bohr in 1924–1926 [8, 9, 10], and many significant contributions were made to the subject in the immediate decade following Bohr's work; see, for example, Bochner [5, 6], von Neumann [30], van Kampen [23]. In the early stage of the theory, much attention was paid to the Fourier series theory of almost periodicity. Later it was observed that many differential equations arising from physics admit almost periodic solutions, then almost periodic phenomenon was extensively studied in differential equations, following Favard's pioneering work [14, 15]; we refer the reader to the books, e.g. Amerio and Prouse [2], Fink [17], Levitan and Zhikov [27], Yoshizawa [37] etc, for the survey.

We know that the white noise perturbations have the effect of mixing and averaging, so what will happen when the almost periodic equation in consideration is perturbed by white noise? In this situation, the almost periodic phenomenon was studied in stochastic differential equations. To the best of our knowledge, only the fixed point method was used so far to investigate the existence of almost periodic solutions by assuming that the linear part of the equation admits the exponential dichotomy; see Halanay [20], Morozan and Tudor [29], Da Prato and Tudor [12], and Arnold and Tudor [3], among others.

In this paper, we aim to adopt the Favard separation method to study almost periodic solutions for stochastic differential equations. The separation method goes back to Favard [14] for linear equations. Consider[1] the linear equation on $\mathbb{R}^d$

$$\dot{x} = A(t)x + f(t). \tag{1.1}$$

If A and f are periodic with common period, the classical Massera criterion [28] states that (1.1) admits a periodic solution with the same period if and only if it admits a bounded solution. When A and f are almost periodic, the situation is more complicated. When A is a constant matrix and f is almost periodic in (1.1), Bohr and Neugebauer [11] proved that a solution of (1.1) is almost periodic if and only if it is bounded. But in the general case, the existence of bounded solutions of (1.1) does not imply the existence of almost periodic ones;

[1]To compare the results for periodic solutions, we reverse somewhat the history of the subject.

see counterexamples given by Zhikov and Levitan [38], Johnson [22], and the more recent work of Ortega and Tarallo [31] which unifies the situations of [38, 22]. Assuming the existence of bounded solutions, Favard [14] proved that (1.1) admits an almost periodic solution if the so-called *Favard separation condition* holds, which means that, for any $B \in H(A)$, each nontrivial bounded solution $x(t)$ of the equation

$$\dot{x} = B(t)x,$$

satisfies

$$\inf_{t \in \mathbb{R}} |x(t)| > 0.$$

Here the hull $H(A)$ of A is defined as follows

$$H(A) = \mathrm{cl}\{A_\tau : \tau \in \mathbb{R}\}$$

with $A_\tau(\cdot) = A(\tau + \cdot)$ and the closure being taken under the uniform topology. The Favard separation condition is optimal in some sense since all the counterexamples we know so far (e.g. the works mentioned above) fail to obey it. The Favard separation condition was extensively studied in the literature in various situations. In particular, Amerio [1] applied Favard's idea to nonlinear differential equations to study almost periodic solutions. Later, Seifert [34] proposed a kind of separation, which is equivalent to almost periodicity, to study the almost periodic solutions of nonlinear equations. Fink [16] generalized separation conditions of [1, 34] to semi-separation ones.

What we mainly concern in the present paper is the existence of almost periodic in distribution solutions to stochastically perturbed differential equations under the Favard or Amerio type separation condition. For instance, when (1.1) is perturbed by small white noise:

$$\mathrm{d}X = (A(t)X + f(t))\mathrm{d}t + \epsilon \mathrm{d}W,$$

does it admit almost periodic solutions in some sense if the unperturbed equation admits bounded solutions and satisfies the Favard separation condition? To this interesting question, the answer is positive. Actually, we can obtain more general result than this; see the following Favard type theorem.

Theorem A. *Consider the Itô stochastic differential equation on $\mathbb{R}^d$*

$$(1.2) \qquad \mathrm{d}X = (A(t)X + f(t))\mathrm{d}t + \sum_{i=1}^{m}(B_i(t)X + g_i(t))\mathrm{d}W_i,$$

where $A, B_1, \ldots, B_m$ are $(d \times d)$-matrix-valued and $f, g_1, \ldots, g_m$ are $\mathbb{R}^d$-valued; all of these functions are almost periodic; $W = (W_1, \ldots, W_m)$ is a standard m-dimensional Brownian motion. Assume that (1.2) admits an L^2-bounded solution X, i.e. $\sup_{t \in \mathbb{R}} E|X(t)|^2 < \infty$, and that the Favard separation condition holds for (1.2). Then (1.2) admits an almost periodic in distribution solution.

For nonlinear stochastic differential equations, we have the following Amerio type theorem.

Theorem B. *Consider the Itô stochastic differential equation on $\mathbb{R}^d$*

$$(1.3) \qquad \mathrm{d}X = f(t, X)\mathrm{d}t + g(t, X)\mathrm{d}W,$$

where $f(t, x)$ is an $\mathbb{R}^d$-valued uniformly almost periodic function, $g(t, x)$ is a $(d \times m)$-matrix-valued uniformly almost periodic function, and W is a standard m-dimensional Brownian motion. Assume that f and g are globally Lipschitz in x with Lipschitz constants independent of t, and that the Amerio semi-separation condition holds for (1.3) in $\mathcal{D}_r$ for some $r > 0$. Then all the L^2-bounded solutions of (1.3), with $\sup_{t \in \mathbb{R}} E|X(t)|^2 \le r^2$, are almost periodic in distribution.

For the Favard separation condition for (1.2) and the Amerio semi-separation condition for (1.3) as well as the meaning of $\mathcal{D}_r$, we refer the reader to Section 2 for details.

Besides the above Theorems A and B, we also obtain a result for linear stochastic equations, which suggests the existence of non-minimal almost periodic in distribution solutions (see Theorem 3.14 for details), a result for nonlinear stochastic equations which weakens, in some sense, the Amerio semi-separation condition in Theorem B (see Theorem 4.6 for details), and a result which reduces the existence of L^2-bounded solutions for (1.3), hence also for (1.2), on the whole real line to that on the positive real line (see Theorem 4.7 for details).

The paper is organized as follows. Section 2 is a preliminary section in which we mainly review some fundamental properties of almost periodic functions and introduce separation conditions for stochastic differential equations. In Section 3, we study almost periodic solutions for linear stochastic equations under the Favard type separation condition. In Section 4, we investigate almost periodic solutions for nonlinear stochastic equations under the Amerio type semi-separation condition. In Section 5, we illustrate our results by some applications. Finally, we discuss, in Section 6, the possibility of improving some results of Sections 3–5, i.e. we can obtain almost periodicity of solutions in distribution sense on the path space.

Throughout the paper, we use $\mathbb{R}$ to denote the set of real numbers, and $\mathbb{R}_- = (-\infty, 0]$, $\mathbb{R}_+ = [0, +\infty)$; we use the same symbol $|\cdot|$ to denote the absolute value of a number, the Euclidian norm of a vector and the induced norm of a matrix, and the cardinality of a set; we denote by B_r the closed ball in $\mathbb{R}^d$ with radius r centered at the origin.

2. Preliminary

Through this section, we assume that (M, d) is a complete metric space.

2.1. Almost periodic functions.

Definition 2.1 (Bohr [8])**.** A continuous function $f : \mathbb{R} \to M$ is called *(Bohr) almost periodic* if for any given $\epsilon > 0$, the set

$$T(\epsilon, f) := \{\tau \in \mathbb{R} : \sup_{t \in \mathbb{R}} d(f(t + \tau), f(t)) < \epsilon\}$$

is *relatively dense* on $\mathbb{R}$, i.e. there is a number $l = l(\epsilon) > 0$ such that $(a, a + l) \cap T(\epsilon, f) \neq \emptyset$ for any $a \in \mathbb{R}$. The set $T(\epsilon, f)$ is called the *set of ϵ-almost periods of f*.

Remark 2.2. For given almost periodic function $f : \mathbb{R} \to M$, it is not hard to prove that f is uniformly continuous on $\mathbb{R}$ and the range $R(f)$ of f is precompact, i.e. the closure of $R(f)$ is compact; see, e.g. [27, page 2].

For simplicity, we follow Bochner's notation [7]. We denote a sequence of real numbers $\{\alpha_n\}$ by α. By $\alpha \subset \beta$ we mean α is a subsequence of β; $-\alpha$ means $\{-\alpha_n\}$; $\alpha > 0$ means $\alpha_n > 0$ for each n; $\{\alpha + \beta\}$ means $\{\alpha_n + \beta_n\}$; α and β being common subsequences of α' and β' means that $\alpha_k = \alpha'_{n(k)}$ and $\beta_k = \beta'_{n(k)}$ for the same function $n(k)$. The notation $T_\alpha f = g$ means $g(t) = \lim_{n \to \infty} f(t + \alpha_n)$ and is written only when the limit exists; the mode of convergence will be specified at each time when the notation is used.

The following definition of almost periodicity is due to Bochner [5].

Definition 2.3. A continuous function $f : \mathbb{R} \to M$ is called *(Bochner) almost periodic* if for any sequence α', there exists a subsequence $\alpha \subset \alpha'$ such that $T_\alpha f$ exists uniformly on $\mathbb{R}$.

Proposition 2.4 (Bochner [5, 7])**.** *For a given continuous function $f : \mathbb{R} \to M$, the following statements are equivalent.*

   (i) *The function f is Bohr almost periodic.*
   (ii) *The function f is Bochner almost periodic.*

(iii) *For every pair of sequences α' and β', there are common subsequences $\alpha \subset \alpha'$ and $\beta \subset \beta'$ such that*

$$T_{\alpha+\beta}f = T_{\alpha}T_{\beta}f \quad pointwise.$$

(iv) *For every pair of sequences α' and β', there are common subsequences $\alpha \subset \alpha'$ and $\beta \subset \beta'$ such that*

$$T_{\alpha+\beta}f = T_{\alpha}T_{\beta}f \quad uniformly\ on\ \mathbb{R}.$$

Remark 2.5. Since Bohr's almost periodicity is equivalent to Bochner's by the above result, we will just call them almost periodicity in what follows.

To study almost periodic solutions of differential equations, we need to consider uniformly almost periodic functions.

Definition 2.6 (Yoshizawa [37])**.** Let $D \subset \mathbb{R}^d$ be an open set. A continuous function $f : \mathbb{R} \times D \to \mathbb{R}^d$ is called *almost periodic in t uniformly for $x \in D$* if for any $\epsilon > 0$ and any compact set $S \subset D$, the set

$$T(\epsilon, f, S) := \{\tau \in \mathbb{R} : \sup_{(t,x)\in\mathbb{R}\times S} |f(t+\tau, x) - f(t,x)| < \epsilon\}$$

is relatively dense on $\mathbb{R}$, i.e. there is a number $l = l(\epsilon, S) > 0$ such that $(a, a+l) \cap T(\epsilon, f, S) \neq \emptyset$ for any $a \in \mathbb{R}$.

Similar to almost periodic functions, we have the following result.

Proposition 2.7 (Yoshizawa [37])**.** *Let $D \subset \mathbb{R}^d$ be an open set. For a given continuous function $f : \mathbb{R} \times D \to \mathbb{R}^d$, the following statements are equivalent.*

(i) *The function f is almost periodic in t uniformly for $x \in D$.*
(ii) *For any sequence α', there exists a subsequence $\alpha \subset \alpha'$ such that $T_{\alpha}f := \lim_{n\to\infty} f(t + \alpha_n, x)$ exists uniformly on $\mathbb{R} \times S$ for any compact $S \subset D$.*
(iii) *For every pair of sequences α' and β', there are common subsequences $\alpha \subset \alpha'$ and $\beta \subset \beta'$ such that for any compact $S \subset D$*

$$T_{\alpha+\beta}f = T_{\alpha}T_{\beta}f \quad uniformly\ on\ \mathbb{R} \times S.$$

For a given function $f : \mathbb{R} \times D \to \mathbb{R}^d$ almost periodic in t uniformly for $x \in D$, the *hull of f* is defined as follows:

$$H(f) := \{g : \text{there exists a sequence } \alpha \text{ such that } T_{\alpha}f = g$$
$$\text{uniformly on } \mathbb{R} \times S \text{ for every compact set } S \subset D\}.$$

We will need the following results in the sequel.

Proposition 2.8. *Let $D \subset \mathbb{R}^d$ be an open set and $f : \mathbb{R} \times D \to \mathbb{R}^d$ be almost periodic in t uniformly for $x \in D$.*

(i) *If a sequence α is such that $T_{\alpha}f$ exists uniformly on $\mathbb{R} \times S$ for any compact $S \subset D$, then $T_{\alpha}f$ is almost periodic in t uniformly for $x \in D$.*
(ii) *Any $g \in H(f)$ is also almost periodic in t uniformly for $x \in D$ and $H(g) = H(f)$.*
(iii) *For any $g \in H(f)$, there exists a sequence α with $\alpha_n \to +\infty$ (or $\alpha_n \to -\infty$) such that $T_{\alpha}f = g$ uniformly on $\mathbb{R} \times S$ for any compact $S \subset \mathbb{R}^d$.*

Remark 2.9. (i) Since we consider stochastic differential equations on $\mathbb{R}^d$ in this paper, i.e. $D = \mathbb{R}^d$ in our situation, we will simply call a function, which is almost periodic in t uniformly for $x \in \mathbb{R}^d$, "uniformly almost periodic" in the sequel if there is no confusion.

(ii) Fink [17] and Seifert [35] introduced slightly different concepts of uniform almost periodicity; see [35] for some discussions on their relations.

2.2. **Asymptotically almost periodic functions.**

Definition 2.10. A continuous function $f : \mathbb{R}_+ \to M$ is *asymptotically almost periodic* if there exists an almost periodic function $p : \mathbb{R} \to M$ such that

$$\lim_{t \to +\infty} d(f(t), p(t)) = 0.$$

The function p is called the *almost periodic part of f*. The asymptotically almost periodic function on $\mathbb{R}_-$ is defined similarly.

Remark 2.11. For a given asymptotically almost periodic function $f : \mathbb{R}_+ \to M$, its almost periodic part is unique.

Proposition 2.12 (Seifert [34], Fink [16])**.** *For a given continuous function $f : \mathbb{R}_+ \to M$, the following statements are equivalent.*

(i) *The function f is asymptotically almost periodic.*
(ii) *For any sequence $\alpha' > 0$ with $\alpha'_n \to +\infty$, there exists a subsequence $\alpha \subset \alpha'$ and a constant $d(\alpha) > 0$ such that $T_\alpha f$ exists pointwise on $\mathbb{R}_+$ and if sequences $\delta > 0$, $\beta \subset \alpha, \gamma \subset \alpha$ are such that*

$$T_{\delta+\beta} f = h_1 \quad and \quad T_{\delta+\gamma} f = h_2$$

exist pointwise on $\mathbb{R}_+$, then either $h_1 \equiv h_2$ on $\mathbb{R}_+$ or $\inf_{t \in \mathbb{R}_+} d(h_1(t), h_2(t)) \geq 2d(\alpha)$.

2.3. **Almost periodicity in distribution.**

Through the paper, we assume for convenience that $(\Omega, \mathcal{F}, P)$ is a probability space which is rich enough to support random variables for any given distribution on $\mathbb{R}^d$ or the path space $C(\mathbb{R}, \mathbb{R}^d)$, the space of $\mathbb{R}^d$-valued continuous functions on $\mathbb{R}$. Let $L^2(P, \mathbb{R}^d)$ stand for the space of all $\mathbb{R}^d$-valued random variables X such that $E|X|^2 = \int_\Omega |X|^2 dP < \infty$. For $X \in L^2(P, \mathbb{R}^d)$, let $\|X\|_2 := \left(\int_\Omega |X|^2 dP \right)^{1/2}$. Then $L^2(P, \mathbb{R}^d)$ is a Hilbert space equipped with the norm $\| \cdot \|_2$. For an $\mathbb{R}^d$-valued stochastic process $X = \{X(t) : t \in \mathbb{R}\}$, if $\sup_{t \in \mathbb{R}} \|X(t)\|_2 < \infty$, we say X is L^2-*bounded* and denote $\|X\|_\infty := \sup_{t \in \mathbb{R}} \|X(t)\|_2$. Then the set of L^2-bounded stochastic processes is a Banach space with the norm $\| \cdot \|_\infty$. In what follows, we also denote by $X(t)$ or $X(\cdot)$ an $\mathbb{R}^d$-valued stochastic process for convenience. Let $\mathcal{P}(\mathbb{R}^d)$ be the space of all Borel probability measures on $\mathbb{R}^d$. For a given $\mathbb{R}^d$-valued random variable X, we denote by $\mathcal{L}(X)$ the law or distribution of X on $\mathbb{R}^d$; for a given process X, by the *law of X on $\mathbb{R}^d$* we mean the $\mathcal{P}(\mathbb{R}^d)$-valued mapping $\mu : \mathbb{R} \to \mathcal{P}(\mathbb{R}^d), t \mapsto \mathcal{L}(X(t))$.

Next, let us introduce the concept of almost periodicity in distribution. For the definiteness, we endow $\mathcal{P}(\mathbb{R}^d)$ with the ρ metric (actually other metrics are also available):

$$\rho(\mu, \nu) := \sup \left\{ \left| \int f d\mu - \int f d\nu \right| : \|f\|_{BL} \leq 1 \right\}, \quad \text{for } \mu, \nu \in \mathcal{P}(\mathbb{R}^d),$$

where f are Lipschitz continuous real-valued functions on $\mathbb{R}^d$ with the norms

$$\|f\|_{BL} = \|f\|_L + \|f\|_\infty, \|f\|_L = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}, \|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|.$$

A sequence $\{\mu_n\} \subset \mathcal{P}(\mathbb{R}^d)$ is said to *weakly converge to μ* if $\int f d\mu_n \to \int f d\mu$ for all $f \in C_b(\mathbb{R}^d)$, the space of all bounded continuous real-valued functions on $\mathbb{R}^d$. It is well-known that $(\mathcal{P}(\mathbb{R}^d), \rho)$ is a separable complete metric space and that a sequence $\{\mu_n\}$ weakly converges to μ if and only if $\rho(\mu_n, \mu) \to 0$ as $n \to \infty$. See [13, Chapter 11] for this metric ρ (denoted by β there) and its related properties. A sequence $\{X_n\}$ of $\mathbb{R}^d$-valued stochastic processes is said to *converge in distribution* to X if $\mathcal{L}(X_n(t))$ weakly converges to $\mathcal{L}(X(t))$; the mode of convergence in t will be specified at each use.

Definition 2.13. An $\mathbb{R}^d$-valued stochastic process X is said to be *(asymptotically) almost periodic in distribution* if its law on $\mathbb{R}^d$ is a $\mathcal{P}(\mathbb{R}^d)$-valued (asymptotically) almost periodic mapping.

Remark 2.14. Since $(\mathcal{P}(\mathbb{R}^d), \rho)$ is a complete metric space, all the assertions on (asymptotic) almost periodicity for the abstract space (M, d) hold for the $\mathbb{R}^d$-valued stochastic processes which are (asymptotically) almost periodic in distribution.

2.4. **Stochastic differential equations and separation.** Assume that W_1 and W_2 are two independent Brownian motions on the probability space $(\Omega, \mathcal{F}, P)$. Let

$$W(t) = \begin{cases} W_1(t), & \text{for } t \geq 0, \\ -W_2(-t), & \text{for } t \leq 0. \end{cases}$$

Then W is a two-sided Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ with $\mathcal{F}_t = \sigma\{W(u) : u \leq t\}, t \in \mathbb{R}$.

Consider the equation (1.3). The triple $(X', W'), (\Omega', \mathcal{F}', P'), \{\mathcal{F}'_t : t \in \mathbb{R}\}$ is a *weak solution* of (1.3) if $(\Omega', \mathcal{F}', P')$ is a probability space and $\{\mathcal{F}'_t : t \in \mathbb{R}\}$ is a filtration of sub-σ-algebras of $\mathcal{F}'$, $W' = \{W'(t) : t \in \mathbb{R}\}$ is an $\mathcal{F}'_t$-adapted m-dimensional Brownian motion and $X' = \{X'(t) : t \in \mathbb{R}\}$ is an $\mathcal{F}'_t$-adapted d-dimensional process such that

$$X'(t) = X'(s) + \int_s^t f(r, X'(r))\mathrm{d}r + \int_s^t g(r, X'(r))\mathrm{d}W'(r)$$

for all $t \geq s$ and each $s \in \mathbb{R}$ almost surely. The weak solution $(X', W'), (\Omega', \mathcal{F}', P'), \{\mathcal{F}'_t : t \in \mathbb{R}\}$ is a *strong solution* if for given $t_0 \in \mathbb{R}$, there exists a measurable function h such that $X'(\cdot) = h(X'(t_0), W'(\cdot))$ on $\mathbb{R}$ almost surely. For strong/weak solutions on the positive real line, see [21] or [24] for details; when the coefficients of (1.3) are globally Lipschitz and of linear growth, see Remark 3.2 for some properties of strong/weak solutions of (1.3) on $\mathbb{R}$.

For the Cauchy problem of (1.3) on the positive real line, it is well-known that the pathwise uniqueness implies uniqueness in the sense of probability law on the path space which we simply call "weak uniqueness", see e.g. [21, §IV.1]; in the meantime, we note that the weak uniqueness implies the uniqueness of law on $\mathbb{R}^d$.

Consider (1.3). To emphasize explicitly the coefficients of (1.3), we also call it *equation (f, g)*. For given $r > 0$, we introduce the following notations:

$$\mathcal{B}_r := \{X \in L^2(P, \mathbb{R}^d) : \|X\|_2 \leq r\}, \quad \mathcal{D}_r := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^2 \mathrm{d}\mu(x) \leq r^2\},$$

$$\mathcal{B}_r^{(1.3)} = \mathcal{B}_r^{(f,g)} := \{X(\cdot) : (X, W) \text{ weakly solves equation } (f, g) \text{ on } \mathbb{R}$$
$$\text{on some filtered probability space for some } W \text{ and } \|X\|_\infty \leq r\},$$

$$\mathcal{D}_r^{(1.3)} = \mathcal{D}_r^{(f,g)} := \{\mu : \mu(\cdot) = \mathcal{L}(X(\cdot)) \text{ for some } X \in \mathcal{B}_r^{(f,g)}\}.$$

Definition 2.15. If $\mathcal{B}_r^{(1.3)}$ is non-empty for some $r > 0$, then $\lambda := \inf_{X \in \mathcal{B}_r^{(1.3)}} \|X\|_\infty$ is called the *minimal value of* (1.3); if $X_0 \in \mathcal{B}_r^{(1.3)}$ and $\|X_0\|_\infty = \lambda$, then X_0 is a *minimal (weak) solution of* (1.3).

Definition 2.16. (i) Assume that the coefficients A, f, B_i, g_i of (1.2) are almost periodic. If there is a sequence α such that $T_\alpha A = \tilde{A}$, $T_\alpha f = \tilde{f}$, $T_\alpha B_i = \tilde{B}_i$ and $T_\alpha g_i = \tilde{g}_i$ uniformly on $\mathbb{R}$ for $i = 1, \ldots, m$, then the equations

$$\mathrm{d}X = (\tilde{A}(t)X + \tilde{f}(t))\mathrm{d}t + \sum_{i=1}^m (\tilde{B}_i(t)X + \tilde{g}_i(t))\mathrm{d}W_i,$$

and

$$\mathrm{d}X = \tilde{A}(t)X\mathrm{d}t + \sum_{i=1}^{m} \tilde{B}_i(t)X\mathrm{d}W_i$$

are called *hull equation of* (1.2) and *homogeneous hull equation of* (1.2), respectively.

(ii) Assume that f, g in (1.3) are uniformly almost periodic. The equation $(\tilde{f}, \tilde{g})$ is called a *hull equation of* (1.3), denoted by $(\tilde{f}, \tilde{g}) \in H(f, g)$, if there exists a sequence α such that $T_\alpha f = \tilde{f}$ and $T_\alpha g = \tilde{g}$, also denoted by $T_\alpha(f, g) = (\tilde{f}, \tilde{g})$, uniformly on $\mathbb{R} \times S$ for any compact subset $S \subset \mathbb{R}^d$.

Definition 2.17. We say that the *Favard (separation) condition holds for* (1.2) if for any homogeneous hull equation corresponding to (1.2)

$$(2.1) \qquad \mathrm{d}X = \tilde{A}(t)X\mathrm{d}t + \sum_{i=1}^{m} \tilde{B}_i(t)X\mathrm{d}W_i,$$

every nontrivial L^2-bounded weak solution X of (2.1) on $\mathbb{R}$ satisfies $\inf_{t \in \mathbb{R}} \|X(t)\|_2 > 0$.

Remark 2.18. Note that if any nontrivial deterministic solution $x(t)$ of the equation $\dot{x} = A(t)x$ satisfies $\inf_{t \in \mathbb{R}} |x(t)| > 0$, then we have $\inf_{t \in \mathbb{R}} \|X(t)\|_2 > 0$ for any nontrivial L^2-bounded stochastic process X which satisfies the same equation. The converse is obviously true. Therefore, the Favard separation condition in Definition 2.17 is a natural generalized version of the usual one mentioned in the Introduction.

Definition 2.19. Assume that f, g in (1.3) are uniformly almost periodic. We say that the *Amerio positive (resp. negative) semi-separation condition holds for* (1.3) in $\mathcal{D}_r$ if any hull equation $(\tilde{f}, \tilde{g})$ of (1.3) only admits positive (resp. negative) semi-separated in distribution solutions in $\mathcal{B}_r$; that is, for any $\mu \in \mathcal{D}_r^{(\tilde{f}, \tilde{g})}$, there is a constant $d(\mu) > 0$, called *separation constant*, such that $\inf_{t \geq 0} \rho(\mu(t), \nu(t)) \geq d(\mu)$ (resp. $\inf_{t \leq 0} \rho(\mu(t), \nu(t)) \geq d(\mu)$) for any other $\nu \in \mathcal{D}_r^{(\tilde{f}, \tilde{g})}$.

Definition 2.20. A property P is called *negative semi-separating in* $\mathcal{D}_r^{(1.3)}$ if for any distinct $\mu, \nu \in \mathcal{D}_r^{(1.3)}$ which satisfy P, there exists a constant $d^{\mu,\nu} > 0$ such that $\inf_{t \in \mathbb{R}_-} \rho(\mu(t), \nu(t)) \geq d^{\mu,\nu}$.

Definition 2.21. Assume that f, g in (1.3) are uniformly almost periodic. A property P is *inherited in distribution in* $\mathcal{D}_r$ if $\mu \in \mathcal{D}_r^{(f,g)}$ has property P with respect to the elements of $\mathcal{D}_r^{(f,g)}$, $(\tilde{f}, \tilde{g}) \in H(f, g)$ with $T_\alpha(f, g) = (\tilde{f}, \tilde{g})$ and $T_\alpha \mu = \nu$ uniformly on compact intervals for some sequence α, then ν also has property P with respect to the elements of $\mathcal{D}_r^{(\tilde{f}, \tilde{g})}$.

3. FAVARD SEPARATION FOR LINEAR STOCHASTIC EQUATIONS

The following result, which simply says that limits of solutions are solutions of the limit equation in distribution sense, is a key "lemma" for what follows and interesting on its own rights, so we state it as a theorem.

Theorem 3.1. *Consider the following family of Itô stochastic equations on* $\mathbb{R}^d$

$$\mathrm{d}X = f_n(t, X)\mathrm{d}t + g_n(t, X)\mathrm{d}W, \quad n = 1, 2, \cdots,$$

where f_n *are* $\mathbb{R}^d$-*valued,* g_n *are* $(d \times m)$-*matrix-valued, and* W *is a standard* m-*dimensional Brownian motion. Assume that* f_n, g_n *satisfy the conditions of global Lipschitz and linear*

growth with common Lipschitz and linear growth constants; that is, there are constants L and K, independent of $t \in \mathbb{R}$ and $n \in \mathbb{N}$, such that for all $x, y \in \mathbb{R}^d$

$$|f_n(t,x) - f_n(t,y)| \vee |g_n(t,x) - g_n(t,y)| \le L|x - y|,$$
$$|f_n(t,x)| \vee |g_n(t,x)| \le K(1 + |x|),$$

where $a \vee b = \max\{a, b\}$. Assume further that $f_n \to f$, $g_n \to g$ pointwise on $\mathbb{R} \times \mathbb{R}^d$ as $n \to \infty$ and that $X_n \in \mathcal{B}_{r_0}^{(f_n, g_n)}$ for some constant r_0, independent of n. Then there is a subsequence of $\{X_n\}$ which converges in distribution, uniformly on compact intervals, to some $X \in \mathcal{B}_{r_0}^{(f,g)}$.

Proof. Fix $n \in \mathbb{N}$. For given bounded interval $[a, b] \subset \mathbb{R}$ and $X_n \in \mathcal{B}_{r_0}^{(f_n, g_n)}$ with (X_n, W_n) weakly solving equation (f_n, g_n) on some filtered probability space, by Cauchy-Schwarz inequality and Itô's isometry we have for any $a \le s \le t \le b$

$$E|X_n(t) - X_n(s)|^2 = E \left| \int_s^t f_n(r, X_n(r))\mathrm{d}r + \int_s^t g_n(r, X_n(r))\mathrm{d}W_n(r) \right|^2$$

$$\le 2(t - s) \int_s^t E|f_n(r, X_n(r))|^2 \mathrm{d}r + 2 \int_s^t E|g_n(r, X_n(r))|^2 \mathrm{d}r$$

$$\le 2(t - s)K^2 \int_s^t E(1 + |X_n(r)|)^2 \mathrm{d}r + 2K^2 \int_s^t E(1 + |X_n(r)|)^2 \mathrm{d}r$$

$$\le 4(t - s)K^2 \int_s^t E(1 + |X_n(r)|^2)\mathrm{d}r + 4K^2 \int_s^t E(1 + |X_n(r)|^2)\mathrm{d}r$$

$$\le 4K^2(\|X_n\|_\infty^2 + 1)(b - a + 1)(t - s)$$

$$(3.1) \qquad \le 4K^2(r_0^2 + 1)(b - a + 1)(t - s).$$

Note that the estimate (3.1) is uniform in $X_n, n \in \mathbb{N}$.

It follows from Chebyshev's inequality that, for any $X \in \mathcal{B}_{r_0}$ and $C \in \mathbb{R}_+$, we have

$$P\{|X| > C\} \le \frac{E|X|^2}{C^2} \le \frac{r_0^2}{C^2}.$$

So for given $\epsilon > 0$, there is a compact set $K_\epsilon := B_C$, a closed ball in $\mathbb{R}^d$, such that

$$P\{X \in \mathbb{R}^d \backslash K_\epsilon\} \le \epsilon$$

by choosing $C > 0$ large enough. By the Prohorov's theorem [33], $\mathcal{D}_{r_0}$ is contained in a compact set in $\mathcal{P}(\mathbb{R}^d)$; actually the Fatou's lemma yields that $\mathcal{D}_{r_0}$ is closed and hence compact in $\mathcal{P}(\mathbb{R}^d)$. It follows from (3.1) that the sequence $\{X_n\}$, regarded as continuous mappings from $[a, b]$ to $L^2(P, \mathbb{R}^d)$, is equi-continuous. Denote $\mu_n(\cdot) = \mathcal{L}(X_n(\cdot)) : [a, b] \to \mathcal{P}(\mathbb{R}^d)$, the law of $X_n(\cdot)$ on $\mathbb{R}^d$. Then since L^2-continuity implies continuity in distribution, the sequence $\{\mu_n\}$ is equi-continuous. Applying a general version of Arzela-Ascoli Theorem (see, e.g. [25, Theorem 7.17]), we obtain a subsequence of $\{\mu_n\}$, still denote by $\{\mu_n\}$, which converges uniformly on $[a, b]$. Since the interval $[a, b]$ is arbitrary, by the diagonal method there is a further subsequence, still denote by $\{\mu_n\}$, such that $\mu_n : \mathbb{R} \to \mathcal{P}(\mathbb{R}^d)$ converges to a function $\mu : \mathbb{R} \to \mathcal{P}(\mathbb{R}^d)$, uniformly on any compact interval.

In the remaining part of the proof, we prove that the limit μ is the law of some L^2-bounded solution X of the equation (f, g) with $\|X\|_\infty \le r_0$, so the theorem is proved.

For the given bounded interval $[a, b]$, since $\mu_n(a) \to \mu(a)$ as $n \to \infty$, by the Skorohod representation theorem there is a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and random variables $\{\tilde{X}_n(a)\}_{n=1}^\infty$, $\tilde{X}(a)$ defined on it so that $\mathcal{L}(\tilde{X}_n(a)) = \mathcal{L}(X_n(a))$, $\mathcal{L}(\tilde{X}(a)) = \mu(a)$ and $\tilde{X}_n(a) \to \tilde{X}(a)$ almost surely as $n \to \infty$. We consider the equation (f_n, g_n) with a common Brownian motion W on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. Then since the coefficients f_n, g_n satisfy the global Lipschitz

and linear growth conditions, by the classical approximation theorem (see, e.g. [19, p54, Theorem 3]), we have

$$(3.2) \qquad \sup_{t \in [a,b]} |\tilde{X}_n(t) - \tilde{X}(t)| \to 0 \quad \text{in probability as } n \to \infty,$$

where $\tilde{X}_n(\cdot)$ and $\tilde{X}(\cdot)$ are strong solutions on $[a, b]$ of equations (f_n, g_n) and (f, g) with the common Brownian motion W and initial values $\tilde{X}_n(a)$ and $\tilde{X}(a)$, respectively. This implies that

$$\mu_n(t) = \mathcal{L}(X_n(t)) = \mathcal{L}(\tilde{X}_n(t)) \to \mathcal{L}(\tilde{X}(t))$$

uniformly on $[a, b]$, where $\mathcal{L}(X_n(t)) = \mathcal{L}(\tilde{X}_n(t))$ holds since the weak uniqueness for equation (f_n, g_n) on $[a, b]$ holds and weak uniqueness implies uniqueness of laws on $\mathbb{R}^d$. On the other hand, $\mu_n(\cdot) \to \mu(\cdot)$ on $[a, b]$. So this enforces that $\mu(t) = \mathcal{L}(\tilde{X}(t))$, $t \in [a, b]$. We may restart from b and repeat the above procedure. In this way, we have proved that μ is the law of some solution of the equation (f, g) on the half line $[a, \infty)$.

Next we will construct, by Kunita's stochastic flow method [26], a strong solution on $(-\infty, a]$ of the equation (f, g) with the above common W so that its law on $\mathbb{R}^d$ coincides with μ on $(-\infty, a]$. Since f and g satisfy the global Lipschitz condition, by [26, Theorem 2.4.3] we know that the solution mapping $\Phi_{s,t}(\cdot, \omega) : \mathbb{R}^d \to \mathbb{R}^d$ of the equation (f, g) is a homeomorphism of $\mathbb{R}^d$ for any $s < t$ and almost all ω. For given $c < a$, we take $\tilde{X}(c, \omega) = \Phi_{c,a}^{-1}(\tilde{X}(a, \omega), \omega)$ for each ω, i.e. the inverse image of $\tilde{X}(a)$ at "time" c. So if we consider the equation (f, g) on $[c, a]$ with initial value $\tilde{X}(c)$, then the value of the solution at "time" a is exactly $\tilde{X}(a)$. In the same way, we take $\tilde{X}_n(c, \omega) = (\Phi_{c,a}^n(\tilde{X}_n(a, \omega), \omega))^{-1}$, with Φ^n being the solution mapping of the equation (f_n, g_n). Then the convergence of $\tilde{X}_n(a)$ to $\tilde{X}(a)$ implies that of $\tilde{X}_n(c)$ to $\tilde{X}(c)$ since Φ^n is a homeomorphism and $\Phi_{c,a}^n \to \Phi_{c,a}$ as $n \to \infty$ by the above mentioned approximation theorem. The same argument as that on $[a, b]$ shows that $\mu(\cdot)$ is the law of $\tilde{X}(\cdot)$ on the interval $[c, a]$. By repeating the procedure, it follows that $\mu(\cdot)$ is the law of $\tilde{X}(\cdot)$ on $(-\infty, a]$ and hence on $\mathbb{R}$.

Since $\tilde{X}_n(t)$ converges in probability to $\tilde{X}(t)$ for each $t \in \mathbb{R}$, the Fatou's lemma and the fact $\|\tilde{X}_n\|_\infty \le r_0$ imply that

$$E|\tilde{X}(t)|^2 \le \liminf_n E|\tilde{X}_n(t)|^2 \le r_0^2.$$

That is, $\|\tilde{X}\|_\infty \le r_0$.

Finally, we replace $\tilde{X}(a)$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ by a random variable $X(a)$ on $(\Omega, \mathcal{F}, P)$ with the same law on $\mathbb{R}^d$, and denote the corresponding solution (the existence and uniqueness is guaranteed by the Lipschitz and linear growth conditions) of equation (f, g) by $X(t)$. Then this solution $X(t)$ admits the same distribution on $\mathbb{R}^d$ as that of $\tilde{X}(t)$ by weak uniqueness for the equation (f, g), and we also have $\|X\|_\infty \le r_0$. This $X(t)$ is what we look for. The proof is complete. $\qquad \square$

Remark 3.2. (i) The above theorem is nontrivial since we consider solutions on the whole real line instead of the usual case where we consider solutions of the Cauchy problem on a positive finite interval or the positive real line.

(ii) For given Brownian motion on some probability space, it follows from the proof of Theorem 3.1 that each solution of (1.3) on $\mathbb{R}$ is determined by the "initial value" at time 0 or at any given "time" $a \in \mathbb{R}$, under the global Lipschitz and linear growth conditions. That is, a weak solution on $\mathbb{R}$ is actually a strong solution on $\mathbb{R}$ in this case; this is similar to the usual case of solutions on the positive real line. Therefore, on some occasions we will not distinguish weak or strong solutions (just call them solutions) in what follows since the equations we consider in this paper satisfy the conditions of global Lipschitz and linear

growth. In this case, for convenience we assume that the probability space $(\Omega, \mathcal{F}, P)$, the Brownian motion W and the filter $\mathcal{F}_t$ are fixed, as pointed out in the Introduction.

(iii) We can also observe from the proof of Theorem 3.1 that if we only consider the law on $\mathbb{R}^d$ of solutions of (1.3) on $\mathbb{R}$, then the law on $\mathbb{R}^d$ is determined by the "initial law" at time 0 or at any given "time" $a \in \mathbb{R}$ by the weak uniqueness on the positive real line and the fact that the Kunita's stochastic flow theorem holds under the global Lipschitz and linear growth conditions. That is, for any given two random variables at 0 or any $a \in \mathbb{R}$ with the same law, the solutions on $\mathbb{R}$ they determine share the same law on $\mathbb{R}^d$. Actually the stronger result holds: they share the same law on the path space.

Lemma 3.3. *Assume that the coefficients f, g of the equation* (1.3) *satisfy the conditions of global Lipschitz and linear growth, and that there is an L^2-bounded solution for* (1.3). *Then* (1.3) *admits a minimal solution.*

Proof. Denote λ as the minimal value of (1.3). Take a sequence $\{X_n\}$ of L^2-bounded solutions of (1.3) such that
$$\|X_n\|_\infty \leq \lambda + \frac{1}{n}.$$
Then it follows from Theorem 3.1 that there is a subsequence of $\{X_n\}$ which converges in distribution to some solution X of (1.3), with $\|X\|_\infty \leq \lambda$. This limit solution is a minimal solution of (1.3). □

Lemma 3.4. *Consider the homogeneous linear equation corresponding to* (1.2) *on $\mathbb{R}^d$*
$$\mathrm{d}X = A(t)X\mathrm{d}t + \sum_{i=1}^m B_i(t)X\mathrm{d}W_i.$$

Assume that A, B_i are almost periodic and that $Y(t)$ is an L^2-bounded solution of the above equation on $\mathbb{R}$ which is almost periodic in distribution. Then we have the following alternative:
$$\inf_{t\in\mathbb{R}} \|Y(t)\|_2 > 0 \quad or \quad Y(t) = 0 \text{ for all } t \in \mathbb{R} \text{ a.s.}$$

Proof. We only need to show that $\inf_{t\in\mathbb{R}} \|Y(t)\|_2 = 0$ implies $Y(t) = 0$ almost surely for all $t \in \mathbb{R}$, which implies $Y(t) = 0$ for all $t \in \mathbb{R}$ almost surely since $Y(t)$ is a continuous process. So let us assume $\inf_{t\in\mathbb{R}} \|Y(t)\|_2 = 0$, then there exists a sequence $\alpha' = \{\alpha'_n\}$ such that $Y(\alpha'_n) \to 0$ in $L^2(P, \mathbb{R}^d)$. It follows from Proposition 2.4 (iv) (by choosing $\beta' = -\alpha'$ there) that there exits a subsequence $\alpha \subset \alpha'$ so that
$$T_\alpha A(t) = \tilde{A}(t), \quad T_\alpha B_i(t) = \tilde{B}_i(t), \quad T_{-\alpha}\tilde{A}(t) = A(t), \quad T_{-\alpha}\tilde{B}_i(t) = B_i(t)$$
and
$$T_\alpha \mu(t) = \tilde{\mu}(t), \quad T_{-\alpha}\tilde{\mu}(t) = \mu(t)$$
uniformly on $\mathbb{R}$, where $\mu(\cdot)$ is the law on $\mathbb{R}^d$ of the solution $Y(\cdot)$. By the proof of Theorem 3.1, the limit $\tilde{\mu}(\cdot)$ is the law of some solution $\tilde{Y}(\cdot)$ of the limit equation
$$\mathrm{d}X = \tilde{A}(t)X\mathrm{d}t + \sum_{i=1}^m \tilde{B}_i(t)X\mathrm{d}W_i.$$

Note that $\tilde{\mu}(0) = T_\alpha\mu(0) = \lim_{n\to\infty} \mu(\alpha_n) = \delta_0$ weakly, with δ_0 being the Dirac measure at 0. So we have $\tilde{Y}(0) = 0$ almost surely, and hence $\tilde{Y}(t) = 0$ almost surely for $t \geq 0$ by the uniqueness of solutions; then by the Kunita's stochastic flow method, as in the proof of Theorem 3.1 again, we have $\tilde{Y}(t) = 0$ almost surely for $t \in \mathbb{R}_-$ and hence $\tilde{\mu}(t) \equiv \delta_0$ for $t \in \mathbb{R}$. So we have $\mu(t) = T_{-\alpha}\tilde{\mu}(t) = \delta_0$ for each $t \in \mathbb{R}$. Therefore, $Y(t) = 0$ almost surely on $\mathbb{R}$. The proof is complete. □

Remark 3.5. Consider the linear equation of the form on $\mathbb{R}^d$

$$(3.3) \qquad \mathrm{d}X = (AX + f(t))\mathrm{d}t + g(t)\mathrm{d}W,$$

where A is a constant matrix, and f, g are almost periodic. Since any non-trivial deterministic bounded solution of $\dot{x} = Ax$ is almost periodic by [17, Theorem 5.3], it follows from Lemma 3.4 that these solutions are separated from 0. So the Favard condition holds for (3.3) by Remark 2.18.

Lemma 3.6. *Assume that the Favard condition holds for the linear equation* (1.2), *then, for given Brownian motion* W *on the filtered probability space* $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$, *there is at most one strong minimal solution for any hull equation of* (1.2).

Proof. If the assertion is not true, then there exists a hull equation

$$(3.4) \qquad \mathrm{d}X = (\tilde{A}(t)X + \tilde{f}(t))\mathrm{d}t + \sum_{i=1}^{m}(\tilde{B}_i(t)X + \tilde{g}_i(t))\mathrm{d}W_i$$

so that X_1 and X_2 are both minimal solutions of the above equation with the common minimal value λ. Then $(X_1 - X_2)/2$ is a nontrivial L^2-bounded solution of the corresponding homogeneous hull equation

$$\mathrm{d}X = \tilde{A}(t)X\mathrm{d}t + \sum_{i=1}^{m}\tilde{B}_i(t)X\mathrm{d}W_i.$$

But the Favard condition yields that there is a constant $\eta > 0$ so that

$$\inf_{t \in \mathbb{R}} \frac{1}{2}\|X_1(t) - X_2(t)\|_2 \geq \eta.$$

It follows from the parallelogram formula that for any $t \in \mathbb{R}$

$$\|\frac{1}{2}(X_1(t) + X_2(t))\|_2^2 + \|\frac{1}{2}(X_1(t) - X_2(t))\|_2^2 = \frac{1}{2}(\|X_1(t)\|_2^2 + \|X_2(t)\|_2^2) \leq \lambda^2.$$

So $\|(X_1 + X_2)/2\|_\infty < \lambda$, the minimal value. This is a contraction since $(X_1 + X_2)/2$ is an L^2-bounded solution of (3.4). $\qquad\square$

Remark 3.7. By Lemma 3.6 and Remark 3.2 (ii)-(iii), it follows that, for any given hull equation of (1.2), all the weak minimal solutions (if they exist) of it share the same law on the path space and hence on $\mathbb{R}^d$ if the Favard condition holds for (1.2).

Lemma 3.8. *Assume that* f, g *in* (1.3) *are uniformly almost periodic and satisfy global Lipschitz condition with Lipschitz constants independent of* t. *Then any hull equation of* (1.3) *admits the same minimal value as that of* (1.3).

Proof. Firstly note that the coefficients f, g of (1.3) satisfy the linear growth condition since $f(\cdot, 0), g(\cdot, 0)$ are bounded on $\mathbb{R}$ by Remark 2.2. Assume that φ is a minimal solution of (1.3), i.e. $\|\varphi\|_\infty = \lambda$, the minimal value of (1.3). Then we have for any $s < t$

$$\varphi(t) = \varphi(s) + \int_s^t f(r, \varphi(r))\mathrm{d}r + \int_s^t g(r, \varphi(r))\mathrm{d}W(r)$$

for some Brownian motion W. Consider the hull equation $(\tilde{f}, \tilde{g})$ with $T_\alpha(f, g) = (\tilde{f}, \tilde{g})$. Denote $\varphi_n(\cdot) := \varphi(\cdot + \alpha_n)$, $f_n(\cdot, \cdot) := f(\cdot + \alpha_n, \cdot)$, $g_n(\cdot, \cdot) := g(\cdot + \alpha_n, \cdot)$, and $W_n(\cdot) := W(\cdot + \alpha_n) - W(\alpha_n)$. Note that f_n, g_n are uniformly almost periodic and globally Lipschitz with the same Lipschitz constants as that of f, g, W_n are standard Brownian motions, and

that $f_n \to \tilde{f}$, $g_n \to \tilde{g}$ uniformly on $\mathbb{R} \times S$ as $n \to \infty$ for any compact subset $S \subset \mathbb{R}^d$. It is clear that φ_n satisfies the following equation for any $s < t$

$$\varphi_n(t) = \varphi_n(s) + \int_s^t f_n(r, \varphi_n(r)) \mathrm{d}r + \int_s^t g_n(r, \varphi_n(r)) \mathrm{d}W_n(r).$$

By Theorem 3.1, there is a subsequence of $\{\varphi_n\}$ which we still denote by the sequence itself so that φ_n converges in distribution, uniformly on compact intervals, to some $\tilde{\varphi}$ as $n \to \infty$ which satisfies the hull equation on $\mathbb{R}$, i.e. for any $s < t$

$$\tilde{\varphi}(t) = \tilde{\varphi}(s) + \int_s^t \tilde{f}(r, \tilde{\varphi}(r)) \mathrm{d}r + \int_s^t \tilde{g}(r, \tilde{\varphi}(r)) \mathrm{d}\tilde{W}(r)$$

for some Brownian motion $\tilde{W}$. The Fatou's lemma implies that $\|\tilde{\varphi}\|_\infty \leq \|\varphi\|_\infty = \lambda$. Hence $\tilde{\lambda} \leq \lambda$, where $\tilde{\lambda}$ denotes the minimal value of the hull equation $(\tilde{f}, \tilde{g})$.

Conversely, by the property of uniform almost periodic functions, we have $T_{-\alpha}\tilde{f} = f$ and $T_{-\alpha}\tilde{g} = g$. By the symmetry, we have $\lambda \leq \tilde{\lambda}$. Therefore, $\lambda = \tilde{\lambda}$. The proof is complete. $\square$

Corollary 3.9. *Consider* (1.3) *and assume that the assumptions of Lemma 3.8 hold. Assume further that φ is a minimal solution of* (1.3)*, and that the sequence α satisfies $T_\alpha(f, g) = (\tilde{f}, \tilde{g})$ and $T_\alpha \varphi$ converges in distribution, uniformly on compact intervals, to some solution $\tilde{\varphi}$ of equation $(\tilde{f}, \tilde{g})$. Then $\tilde{\varphi}$ is a minimal solution of equation $(\tilde{f}, \tilde{g})$.*

Proof. It is immediate from the proof of Lemma 3.8. $\square$

Lemma 3.10. *Assume that each hull equation of* (1.3) *admits a unique minimal solution in distribution sense, i.e. all the minimal solutions of the given hull equation possess the same law on $\mathbb{R}^d$. Then these minimal solutions are almost periodic in distribution.*

Proof. For given $(\tilde{f}, \tilde{g}) \in H(f, g)$ and arbitrary sequences α' and β', by the property of uniform almost periodic functions there exist common subsequences α, β of α', β' so that

$$T_{\alpha+\beta}\tilde{f} = T_\alpha T_\beta \tilde{f}, \quad T_{\alpha+\beta}\tilde{g} = T_\alpha T_\beta \tilde{g}$$

uniformly on $\mathbb{R} \times S$ for any compact subset S of $\mathbb{R}^d$. Assume that $\tilde{\varphi}$ is a minimal solution of the equation $(\tilde{f}, \tilde{g})$ whose law on $\mathbb{R}^d$ is $\tilde{\mu}$.

By the proof of Theorem 3.1, there exist common subsequences of α and β, which we still denote by α and β, so that $T_{\alpha+\beta}\tilde{\mu}$ and $T_\alpha T_\beta \tilde{\mu}$ exist, uniformly on compact intervals, and they are laws of processes φ_1 and φ_2, which are solutions of the equations $(T_{\alpha+\beta}\tilde{f}, T_{\alpha+\beta}\tilde{g})$ and $(T_\alpha T_\beta \tilde{f}, T_\alpha T_\beta \tilde{g})$, respectively. That is, φ_1 and φ_2 satisfy the same equation. By Corollary 3.9, both φ_1 and φ_2 are minimal solutions of the equation $(T_{\alpha+\beta}\tilde{f}, T_{\alpha+\beta}\tilde{g})$. But each hull equation admits a unique minimal solution in distribution sense, which enforces $\mathcal{L}(\varphi_1) = \mathcal{L}(\varphi_2)$ and hence $T_{\alpha+\beta}\tilde{\mu} = T_\alpha T_\beta \tilde{\mu}$. That is, $\tilde{\varphi}$ is almost periodic in distribution by Proposition 2.4. The proof is complete. $\square$

The following result is Theorem A in the Introduction.

Theorem 3.11. *Consider* (1.2) *with the coefficients $A, B_1, \ldots, B_m$, $f, g_1, \ldots, g_m$ being almost periodic. Assume further that* (1.2) *admits an L^2-bounded solution and that the Favard condition holds for* (1.2)*. Then* (1.2) *admits an almost periodic in distribution solution.*

Proof. It follows from Lemma 3.3 and Corollary 3.9 that each hull equation of (1.2) admits minimal solutions, and by Lemma 3.6 and Remark 3.7 each hull equation admits a unique minimal solution in distribution sense. So the theorem follows from Lemma 3.10. $\square$

Remark 3.12. It follows from Remark 3.5 and Theorem 3.11 that the existence of L^2-bounded solutions of (3.3) implies that it admits an almost periodic in distribution solution. This

can be regarded as a stochastic version of Bohr-Neugebauer type result, mentioned in the Introduction.

Corollary 3.13. *Consider the equation of the form on $\mathbb{R}^d$*

$$(3.5) \qquad \mathrm{d}X = [A(t)X + f(t)]\mathrm{d}t + g(t)\mathrm{d}W,$$

where A, f, g are almost periodic functions. If the corresponding deterministic equation

$$(3.6) \qquad \dot{x} = A(t)x + f(t)$$

satisfies the Favard condition in usual sense (mentioned in the Introduction) and (3.5) admits an L^2-bounded solution, then (3.5) admits an almost periodic in distribution solution.

Proof. Note that the homogeneous equation corresponding to (3.5) is the same as that of the deterministic equation (3.6), so the Favard condition holds for (3.5) by Remark 2.18. The result now follows from Theorem 3.11. □

Finally we give a result, which confirms that there may be other almost periodic in distribution solutions besides minimal ones.

Theorem 3.14. *Consider the linear equation on $\mathbb{R}^d$*

$$(3.7) \qquad \mathrm{d}X = (AX + f(t))\mathrm{d}t + g(t)\mathrm{d}W,$$

where A is a constant matrix, f, g are almost periodic, and W is a given m-dimensional Brownian motion. If X is a strong L^2-bounded solution of (3.7) on $\mathbb{R}$ so that $X(\tau) - X_0(\tau)$ is independent of $X_0(\tau)$ and W for some $\tau \in \mathbb{R}$, where X_0 is the strong minimal solution of (3.7). Then X is almost periodic in distribution.

Proof. By Remark 3.5 we know that the Favard condition holds for (3.7), so it follows from Lemmas 3.3 and 3.6 that there is a unique strong minimal solution X_0 for equation (3.7), which is almost periodic in distribution by Lemma 3.10. Let $Y(\cdot) = X(\cdot) - X_0(\cdot)$. Then Y is an L^2-bounded solution of the equation $\dot{x} = Ax$ on $\mathbb{R}$, which is almost periodic in L^2-sense by [17, Theorem 5.3][2], i.e. the mapping $t \mapsto Y(t)$ is an almost periodic $L^2(P, \mathbb{R}^d)$-valued mapping. By the Bochner's definition for almost periodicity, it follows that $Y(\cdot)$ is almost periodic in distribution since L^2-convergence implies convergence in distribution.

Since X_0 is the solution of (3.7) with "initial value" $X_0(\tau)$, by the strong solution theorem of Yamada-Watanabe (see [36] or [21, Theorem IV.1.1]),

$$(3.8) \qquad X_0(\cdot) = F(X_0(\tau), W) \text{ for some measurable function } F$$

and $t \geq \tau$. Since the coefficients of (3.7) are globally Lipschitz in x, Kunita's stochastic flow theorem implies that we may regard that (3.8) holds for all $t \in \mathbb{R}$. Similarly, for the equation $\dot{x} = Ax$, we have $Y(\cdot) = G(Y(\tau))$ for some measurable function G. Since the random variable $Y(\tau)$ is independent of $X_0(\tau)$ and W, $Y(\cdot)$ is independent of $X_0(\cdot)$. In particular, $X_0(t)$ is independent of $Y(t)$ for each $t \in \mathbb{R}$. So we have

$$\mathcal{L}(X(t)) = \mathcal{L}(X_0(t)) * \mathcal{L}(Y(t)) \qquad \text{for each } t \in \mathbb{R}.$$

Denote $\mu_1(t) = \mathcal{L}(X_0(t))$ and $\mu_2(t) = \mathcal{L}(Y(t))$ for each $t \in \mathbb{R}$. For arbitrary sequences α' and β', it follows from the almost periodicity of μ_1 and μ_2 that there are common subsequences $\alpha \subset \alpha'$ and $\beta \subset \beta'$ such that

$$T_{\alpha+\beta}\mu_i(t) = T_\alpha T_\beta \mu_i(t) \qquad \text{for each } t \in \mathbb{R}, i = 1, 2.$$

Since the convolution of probability measures is continuous (see, e.g. [13, Theorem 9.5.9]), it follows that

$$T_{\alpha+\beta}[\mu_1(t) * \mu_2(t)] = T_\alpha T_\beta[\mu_1(t) * \mu_2(t)] \qquad \text{for each } t \in \mathbb{R}.$$

---

[2] Note that the argument there still applies when the initial value is replaced by an L^2-random variable.

That is, $\mathcal{L}(X(\cdot))$ is almost periodic by Proposition 2.4. The proof is complete.   □

4. Amerio separation for nonlinear stochastic equations

In this section, we consider the nonlinear equation (1.3). Firstly, let us state the following standing hypothesis which is used frequently in the sequel:

(H) Assume that $f(t, x)$ is an $\mathbb{R}^d$-valued uniformly almost periodic function, $g(t, x)$ is a $(d \times m)$-matrix-valued uniformly almost periodic function, and W is a standard m-dimensional Brownian motion. Assume further that f and g are globally Lipschitz in x with Lipschitz constants independent of t.

Lemma 4.1. *Consider* (1.3) *and assume* **(H)**. *If* (1.3) *admits an L^2-bounded solution X on $\mathbb{R}$ which is asymptotically almost periodic in distribution on $\mathbb{R}_+$, then* (1.3) *admits a solution Y on $\mathbb{R}$ which is almost periodic in distribution such that*

$$\lim_{t \to +\infty} \rho(\mathcal{L}(X(t)), \mathcal{L}(Y(t))) = 0 \quad and \quad \|Y\|_\infty \le \|X\|_\infty.$$

In particular, $\mathcal{L}(Y)$ is the almost periodic part of $\mathcal{L}(X)$. The similar result holds when X is asymptotically almost periodic in distribution on $\mathbb{R}_-$.

Proof. Denote $\mu(t) := \mathcal{L}(X(t))$ for each $t \in \mathbb{R}$ and fix a sequence $\alpha' \subset \mathbb{R}_+$ with $\alpha'_n \to \infty$. Since f, g are uniformly almost periodic, there is a subsequence α of α' such that $T_\alpha f$ and $T_\alpha g$ uniformly exist on every $\mathbb{R} \times S$, with $S \subset \mathbb{R}^d$ being compact. By the proof of Theorem 3.1, the above subsequence α can be chosen such that $T_\alpha \mu$ exists uniformly on any compact interval of $\mathbb{R}$. On the other hand, since μ is asymptotically almost periodic on $\mathbb{R}_+$, there is a $\mathcal{P}(\mathbb{R}^d)$-valued almost periodic function η such that $\lim_{t \to \infty} \rho(\mu(t), \eta(t)) = 0$. Note that the above subsequence α can be chosen such that $T_\alpha \eta$ uniformly exits on $\mathbb{R}$ by the almost periodicity of η (so $T_\alpha \eta$ is almost periodic by Proposition 2.8 (i)) and

$$T_\alpha \mu(t) = T_\alpha \eta(t) \quad \text{for all } t \in \mathbb{R}.$$

Also the proof of Theorem 3.1 implies that $T_\alpha \mu$ (and hence $T_\alpha \eta$) is the law of some L^2-bounded solution $\tilde{X}(t)$, with $\|\tilde{X}\|_\infty \le \|X\|_\infty$, of the limit equation

$$\mathrm{d}X(t) = T_\alpha f(t, X(t))\mathrm{d}t + T_\alpha g(t, X(t))\mathrm{d}W(t).$$

We take a subsequence of α if necessary (still denote it by α) such that

$$T_{-\alpha} T_\alpha f = f, \ T_{-\alpha} T_\alpha g = g$$

uniformly on every $\mathbb{R} \times S$ and

$$T_{-\alpha} T_\alpha \eta = \eta$$

uniformly on $\mathbb{R}$. It follows from the proof of Theorem 3.1 again that $T_{-\alpha} T_\alpha \eta$ is the law of some L^2-bounded solution Y of the equation

$$\mathrm{d}X(t) = T_{-\alpha} T_\alpha f(t, X(t))\mathrm{d}t + T_{-\alpha} T_\alpha g(t, X(t))\mathrm{d}W(t)$$

with $\|Y\|_\infty \le \|\tilde{X}\|_\infty$. That is, the almost periodic function η is the law of the solution Y of the equation (1.3).

The proof in the case of X being asymptotically almost periodic on $\mathbb{R}_-$ is similar.   □

The following Amerio type result is Theorem B in the Introduction.

Theorem 4.2. *Consider* (1.3). *Assume* **(H)** *and that the Amerio positive (or negative) semi-separation condition holds for* (1.3) *in $\mathcal{D}_r$. Then $|\mathcal{D}_r^{(1.3)}|$ is finite. If $\mathcal{B}_r^{(1.3)}$ is non-empty, then it consists of almost periodic in distribution solutions of* (1.3).

Proof. We only consider the case of positive semi-separation since the negative semi-separation case is similar.

Firstly, $\mathcal{D}_r^{(1.3)}$ consists of finite number of elements, so the separation constant for (1.3) depends only on the equation itself, independent of elements of $\mathcal{D}_r^{(1.3)}$. Indeed, by Theorem 3.1, if there are infinite elements in $\mathcal{D}_r^{(1.3)}$, then there exists a subsequence which converges, uniformly on compact intervals, to an element of $\mathcal{D}_r^{(1.3)}$. But this limit cannot be positively semi-separated, a contradiction to the Amerio semi-separation condition.

Secondly, the separation constant can be taken the same for all the hull equations of (1.3). To see this, for any $(\tilde{f}, \tilde{g}) \in H(f, g)$, by Proposition 2.8 (iii) there exists a sequence α' with $\alpha'_n \to +\infty$ such that $T_{\alpha'}(f, g) = (\tilde{f}, \tilde{g})$ uniformly on $\mathbb{R} \times S$ for any compact $S \subset \mathbb{R}^d$. Denote $\mathcal{D}_r^{(1.3)} = \{\mu_1, \ldots, \mu_\kappa\}$ and the separation constant for (1.3) by $d^{(f,g)}$. Then there exists a subsequence α of α' such that $T_\alpha \mu_i(\cdot), i = 1, \ldots, \kappa$ exist uniformly on any compact interval and $T_\alpha \mu_i \in \mathcal{D}_r^{(\tilde{f}, \tilde{g})}$ by the proof of Theorem 3.1. Note that $\inf_{t \in \mathbb{R}_+} \rho(\mu_i(t), \mu_j(t)) \geq d^{(f,g)}$ implies that

$$(4.1) \qquad \inf_{t \in \mathbb{R}_+} \rho(T_\alpha \mu_i(t), T_\alpha \mu_j(t)) \geq d^{(f,g)} \qquad \text{for } 1 \leq i, j \leq \kappa \text{ and } i \neq j.$$

That is, $T_\alpha \mu_i$ are distinct and so $|\mathcal{D}_r^{(\tilde{f}, \tilde{g})}| \geq \kappa$. Conversely, it follows from the fact $(f, g) \in H(\tilde{f}, \tilde{g})$ that the above argument holds symmetrically. So $|\mathcal{D}_r^{(f,g)}|$ is no less than $|\mathcal{D}_r^{(\tilde{f}, \tilde{g})}|$. This enforces that $|\mathcal{D}_r^{(\tilde{f}, \tilde{g})}| = |\mathcal{D}_r^{(f,g)}|$. In the meantime, it follows from (4.1) that $\mathcal{D}_r^{(\tilde{f}, \tilde{g})} = \{T_\alpha \mu_1, \ldots, T_\alpha \mu_\kappa\}$ and $d^{(\tilde{f}, \tilde{g})} \geq d^{(f,g)}$, with $d^{(\tilde{f}, \tilde{g})}$ being the separation constant for the equation $(\tilde{f}, \tilde{g})$. By symmetry, we have $d^{(f,g)} \geq d^{(\tilde{f}, \tilde{g})}$. So $d^{(f,g)} = d^{(\tilde{f}, \tilde{g})}$, which we denote by $d^{H(f,g)}$.

Thirdly, any $X \in \mathcal{B}_r^{(1.3)}$, with $\mu(\cdot) = \mathcal{L}(X(\cdot))$, is asymptotically almost periodic in distribution on $\mathbb{R}_+$. Indeed, for any sequence $\eta' > 0$ with $\eta'_n \to \infty$, there exists a subsequence $\eta \subset \eta'$ such that $T_\eta(f, g)$ exists uniformly on $\mathbb{R} \times S$ for any compact $S \subset \mathbb{R}^d$ by the uniform almost periodicity of f, g and $T_\eta \mu$ exists uniformly on any compact interval by the proof of Theorem 3.1. Assume that sequences $\delta > 0$, $\beta \subset \eta, \gamma \subset \eta$ are such that

$$T_{\delta+\beta}\mu = \nu_1 \quad \text{and} \quad T_{\delta+\gamma}\mu = \nu_2$$

exist pointwise on $\mathbb{R}_+$. By taking subsequences of δ, β, γ if necessary, we may assume that $T_{\delta+\beta}\mu, T_{\delta+\gamma}\mu$ exist uniformly on any compact interval of $\mathbb{R}$, and

$$T_{\delta+\beta}(f, g) = T_\delta T_\beta(f, g) = T_\delta T_\eta(f, g), \quad T_{\delta+\gamma}(f, g) = T_\delta T_\gamma(f, g) = T_\delta T_\eta(f, g)$$

uniformly on $\mathbb{R} \times S$ for any compact S. So by the proof of Theorem 3.1 again, ν_1 and ν_2 are the laws on $\mathbb{R}^d$ of solutions for the same equation $T_\delta T_\eta(f, g)$. Then it follows from the Amerio positive separation condition that $\nu_1(t) \equiv \nu_2(t)$ or $\rho(\nu_1(t), \nu_2(t)) \geq d^{H(f,g)}$ on $\mathbb{R}_+$. That is, μ is an asymptotically almost periodic $\mathcal{P}(\mathbb{R}^d)$-valued mapping by Proposition 2.12. Therefore, X is asymptotically almost periodic in distribution on $\mathbb{R}_+$.

Finally, the above given $X \in \mathcal{B}_r^{(1.3)}$ is actually almost periodic in distribution. To see this, note that the almost periodic part p of the law $\mu(\cdot) = \mathcal{L}(X(\cdot))$ is indeed the law of some solution of (1.3) by Lemma 4.1. That is, $\mu, p \in \mathcal{D}_r^{(1.3)}$. But $\lim_{t \to +\infty} \rho(\mu(t), p(t)) = 0$, so the Amerio semi-separation condition enforces that $\mu(t) \equiv p(t)$ on $\mathbb{R}$. That is, X is almost periodic in distribution. The proof is complete. $\square$

Remark 4.3. Note that, under the assumptions of Theorem 4.2, we have $|\mathcal{D}_r^{(\tilde{f}, \tilde{g})}| = |\mathcal{D}_r^{(f,g)}|$ for any hull equation $(\tilde{f}, \tilde{g})$ of (1.3) and all the elements of $\mathcal{D}_r^{(\tilde{f}, \tilde{g})}$ are almost periodic.

The following trivial separation case is very important in applications.

Corollary 4.4. *Consider* (1.3). *Assume* **(H)** *and that each hull equation* $(\tilde{f}, \tilde{g})$ *of* (1.3) *admits a unique distribution in* $\mathcal{D}_r$, *i.e.* $|\mathcal{D}_r^{(\tilde{f},\tilde{g})}| = 1$. *Then* $\mathcal{B}_r^{(1.3)}$ *consists of almost periodic in distribution solutions with the unique common distribution in* $\mathcal{D}_r^{(1.3)}$.

Remark 4.5. In the literature, the only applications of deterministic Amerio type theorems to specific models are the trivial separation case, i.e. there is a unique solution in some given compact subset of $\mathbb{R}^d$ ($\mathcal{D}_r$ in our case); see, e.g. [17]. For stochastic equations, it is certainly interesting to find (if possible) applications in nontrivial separation case.

One weakness of Amerio type results (Theorem 4.2 and Corollary 4.4) is that they impose hypotheses on all hull equations. This may be partly remedied by the inheritance property, as stated in the following result.

Theorem 4.6. *Consider* (1.3) *and assume* **(H)**. *Assume that the property P is negative semi-separating in* $\mathcal{D}_r^{(1.3)}$ *and inherited in distribution in* $\mathcal{D}_r$, *and that the number of elements of* $\mathcal{D}_r^{(1.3)}$ *satisfying property P is finite. Then every element of* $\mathcal{D}_r^{(1.3)}$ *with property P is almost periodic. In particular,* (1.3) *admits almost periodic in distribution solutions in* $\mathcal{B}_r$.

Proof. Assume that the elements of $\mathcal{D}_r^{(1.3)}$ with property P are $\mu_1, \ldots, \mu_\kappa$, so the separation constant depends only on the equation (1.3), which we denote by $d^{(f,g)}$. That is, $\inf_{t \in \mathbb{R}_-} \rho(\mu_i(t), \mu_j(t)) \geq d^{(f,g)}$ for $i, j = 1, \ldots, \kappa$ and $i \neq j$.

We now show that any hull equation $(\tilde{f}, \tilde{g})$ of (1.3) also admits κ elements of $\mathcal{D}_r^{(\tilde{f},\tilde{g})}$ with property P and the separation constant for the equation $(\tilde{f}, \tilde{g})$ can be chosen the same as that of (1.3). Indeed, by Proposition 2.8 (iii) we may take a sequence α' with $\alpha'_n \to -\infty$ such that $T_{\alpha'}(f, g) = (\tilde{f}, \tilde{g})$ uniformly on $\mathbb{R} \times S$ for any compact $S \subset \mathbb{R}^d$. Take a subsequence α of α' such that $T_\alpha \mu_i(\cdot), i = 1, \ldots, \kappa$ exist uniformly on any compact interval and $T_\alpha \mu_i \in \mathcal{D}_r^{(\tilde{f},\tilde{g})}$ by the proof of Theorem 3.1. Furthermore, $\inf_{t \in \mathbb{R}_-} \rho(T_\alpha \mu_i(t), T_\alpha \mu_j(t)) \geq d^{(f,g)}$ for $i \neq j$ and it follows from the inheritance of property P that each $T_\alpha \mu_i$ also satisfies property P. That is, the equation $(\tilde{f}, \tilde{g})$ admits at least κ elements of $\mathcal{D}_r^{(\tilde{f},\tilde{g})}$ with property P. By symmetry, there exists a sequence β' with $\beta'_n \to -\infty$ such that $T_{\beta'}(\tilde{f}, \tilde{g}) = (f, g)$, so the same argument yields that equation $(\tilde{f}, \tilde{g})$ admits exactly κ elements of $\mathcal{D}_r^{(\tilde{f},\tilde{g})}$ with property P which are $T_\alpha \mu_i, i = 1, \ldots, \kappa$. The symmetric argument also implies that the separation constant among these elements of $\mathcal{D}_r^{(\tilde{f},\tilde{g})}$ with property P can be chosen the same as that of (1.3), which we denote by $d^{H(f,g)}$.

The remaining proof is almost the same as that of Theorem 4.2, just replacing $\mathbb{R}_+$ by $\mathbb{R}_-$. The proof is complete. $\square$

In applications, the inheritance property is usually checked through stability properties. Since this topic itself deserves a separate paper, here we will not discuss it further in this direction.

To conclude this section, we give a sufficient condition on the existence of L^2-bounded solutions.

Theorem 4.7. *Consider* (1.3) *and assume* **(H)**. *If* (1.3) *admits a solution φ on $[t_0, \infty)$ with* $\sup_{t \in [t_0,\infty)} \|\varphi(t)\|_2 \leq M$ *for some $t_0 \in \mathbb{R}$ and constant $M > 0$, then* (1.3) *admits an L^2-bounded solution $\tilde{\varphi}$ on $\mathbb{R}$ satisfying* $\|\tilde{\varphi}\|_\infty \leq M$.

Proof. By Proposition 2.7 (i)–(ii) and Proposition 2.8 (iii), we may take a sequence α such that $T_\alpha(f, g) = (f, g)$ uniformly on $\mathbb{R} \times S$ for any compact $S \subset \mathbb{R}^d$, with $\alpha_n \to +\infty$ as $n \to \infty$.

Note that for any $t \geq s \geq t_0$

$$\varphi(t) = \varphi(s) + \int_s^t f(r, \varphi(r))\mathrm{d}r + \int_s^t g(r, \varphi(r))\mathrm{d}W(r)$$

for some Brownian motion W. Similar to the proof of Lemma 3.8, let φ_n, f_n, g_n, W_n be defined the same as there. Then φ_n is defined on $[t_0 - \alpha_n, \infty)$, satisfies $\sup_{r \in [t_0 - \alpha_n, \infty)} \|\varphi_n(r)\|_2 \leq M$ and

$$\varphi_n(t) = \varphi_n(s) + \int_s^t f_n(r, \varphi_n(r))\mathrm{d}r + \int_s^t g_n(r, \varphi_n(r))\mathrm{d}W_n(r)$$

for any $t \geq s \geq t_0 - \alpha_n$.

For any fixed $a \in \mathbb{R}$, φ_n is defined on $[a, \infty)$ when n is large, and by Theorem 3.1 there is a subsequence of $\{\varphi_n\}$ which converges in distribution, uniformly on compact intervals of $[a, \infty)$, to some $\tilde{\varphi}$ as $n \to \infty$ which satisfies the limit equation (i.e. (1.3) itself) on $[a, \infty)$, i.e. for any $a \leq s \leq t$

$$\tilde{\varphi}(t) = \tilde{\varphi}(s) + \int_s^t f(r, \tilde{\varphi}(r))\mathrm{d}r + \int_s^t g(r, \tilde{\varphi}(r))\mathrm{d}\tilde{W}(r)$$

for some Brownian motion $\tilde{W}$. By choosing a to be a sequence converging to $-\infty$ and the standard diagonal method, we may assume that the limit $\tilde{\varphi}$ satisfies (1.3) on $\mathbb{R}$. The Fatou's lemma implies that $\|\tilde{\varphi}\|_\infty \leq \sup_{t \in [t_0, \infty)} \|\varphi(t)\|_2 \leq M$. The proof is complete. $\qquad\square$

5. Applications

In this section, we give some applications of our results.

Theorem 5.1. *Consider* (1.3) *and assume* **(H)**. *Assume that the function* $V : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}_+$ *is* C^2 *in* $t \in \mathbb{R}$ *and* C^3 *in* $x \in \mathbb{R}^d$, *and that the differentials* $D^i V$ *of* V *with* $i = 0, 1, 2$ *and the derivatives* $V_{tx_ix_j}$, $V_{x_ix_jx_k}$, $i, j, k = 1, \ldots, d$ *are bounded on* $\mathbb{R} \times S$ *for any compact* $S \subset \mathbb{R}^d$. *Assume further that* V *satisfies:*

$$\mathscr{L}V(t, x - y) := \frac{\partial V}{\partial t}(t, x - y) + \sum_{i=1}^d (f_i(t, x) - f_i(t, y))\frac{\partial V}{\partial x_i}(t, x - y)$$

$$+ \frac{1}{2} \sum_{i,j=1}^d \sum_{l=1}^m (g_{il}(t, x) - g_{il}(t, y))\frac{\partial^2 V}{\partial x_i \partial x_j}(t, x - y)(g_{jl}(t, x) - g_{jl}(t, y))$$

(h_0) $\qquad\qquad \geq a(|x - y|) \quad$ *for all* $t \in \mathbb{R}$ *and* $x, y \in \mathbb{R}^d$,

(h_1) $\qquad\qquad \inf_{t \in \mathbb{R}} V(t, x) > 0$ *for each* $x \neq 0$, *and* $V(t, 0) = 0$ *for any* $t \in \mathbb{R}$,

and

(h_2) $\qquad V(t, x) \leq b|x|^2 + c$ *for all* $(t, x) \in \mathbb{R} \times \mathbb{R}^d$ *and some positive constants* b, c,

where $a : \mathbb{R}_+ \to \mathbb{R}_+$ *is continuous,* $a(0) = 0$, $a(r) > 0$ *for* $r > 0$ *and* $\liminf_{r \to \infty} a(r) > 0$. *Then* $\cup_{r>0} \mathcal{D}_r^{(1.3)}$ *is empty or consists of a unique element which is almost periodic in* t.

Proof. We divide the proof into 3 steps.

Step 1. Uniqueness of strong L^2-bounded solutions. Assume that X and Y are two strong L^2-bounded solutions of (1.3) on $\mathbb{R}$ for given Brownian motion W. Fix $t_0 \in \mathbb{R}$. Then for any $t \geq t_0$,

$$X(t) - Y(t) = X(t_0) - Y(t_0) + \int_{t_0}^t [f(s, X(s)) - f(s, Y(s))]\mathrm{d}s$$

$$+ \int_{t_0}^{t} [g(s, X(s)) - g(s, Y(s))] \mathrm{d}W(s).$$

We define a sequence of stopping times as follows:

$$\tau_n := \inf\{t \geq t_0 : \max\{|X(t)|, |Y(t)|\} \geq n\}.$$

Then we have by Itô's formula and (h_0)

$$EV(t \wedge \tau_n, X(t \wedge \tau_n) - Y(t \wedge \tau_n)) = EV(t_0, X(t_0) - Y(t_0)) + E \int_{t_0}^{t \wedge \tau_n} \mathscr{L}V(s, X(s) - Y(s)) \mathrm{d}s$$

$$\geq EV(t_0, X(t_0) - Y(t_0)) + E \int_{t_0}^{t \wedge \tau_n} a(|X(s) - Y(s)|) \mathrm{d}s.$$

Noting that $\tau_n \to \infty$ almost surely as $n \to \infty$, we have

(5.1) $$EV(t, X(t) - Y(t)) \geq EV(t_0, X(t_0) - Y(t_0)) + E \int_{t_0}^{t} a(|X(s) - Y(s)|) \mathrm{d}s.$$

In particular,

(5.2) $$EV(t, X(t) - Y(t)) \geq EV(t_0, X(t_0) - Y(t_0)).$$

By the L^2-boundedness of $X(t)$, $Y(t)$ and (h_2), the limit $\lim_{t \to \infty} EV(t, X(t) - Y(t))$ exists. This together with (5.1) implies that

$$\lim_{n \to \infty} E \int_{n}^{\infty} a(|X(s) - Y(s)|) \mathrm{d}s = 0.$$

Since a is nonnegative, there exists a sequence $\{t_n\}$, with $t_n \to \infty$ as $n \to \infty$, such that $\lim_{n \to \infty} Ea(|X(t_n) - Y(t_n)|) = 0$. Since the function a only vanishes at 0 and $\liminf_{r \to \infty} a(r) > 0$, this enforces that $\lim_{n \to \infty} |X(t_n) - Y(t_n)| = 0$ in probability. So the Lebesgue dominated convergence theorem and (h_1)-(h_2) yield that

$$\lim_{n \to \infty} EV(t_n, X(t_n) - Y(t_n)) = 0,$$

which implies $EV(t_0, X(t_0) - Y(t_0)) = 0$ by (5.2) and the non-negativeness of V. So $X(t_0) = Y(t_0)$ almost surely by (h_1) again. Since f, g are global Lipschitz, we have $X(t) = Y(t)$ on $\mathbb{R}$ almost surely by Remark 3.2 (ii).

Step 2. Convergence of V and inheritance of (h_0)-(h_2). For given sequence α', let $V_n(\cdot, \cdot) := V(\cdot + \alpha'_n, \cdot)$. For any compact interval $I \subset \mathbb{R}$ and compact subset $S \subset \mathbb{R}^d$, since V, V_t, V_{x_i} are bounded on $\mathbb{R} \times S$, V_n are uniformly bounded and equi-continuous on $I \times S$. So it follows from the Arzela-Ascoli theorem that there exists a subsequence α of α' so that $T_\alpha V = \lim_{n \to \infty} V(\cdot + \alpha_n, \cdot)$ uniformly exists on $I \times S$; by the diagonalization argument, the subsequence α may be chosen such that $T_\alpha V$ uniformly exists on any compact subsets of $\mathbb{R} \times \mathbb{R}^d$. In the same way, by the hypothesis on the bounded differentials and derivatives of V, the subsequence α can be further chosen such that $T_\alpha V_t, T_\alpha V_{x_i}, T_\alpha V_{x_i x_j}$ (the meaning of these notations is like $T_\alpha V$) uniformly exists on any compact subsets of $\mathbb{R} \times \mathbb{R}^d$.

Since $T_\alpha V$ and $T_\alpha V_t$ uniformly exist on any compact subsets of $\mathbb{R} \times \mathbb{R}^d$, it follows that $\frac{\partial T_\alpha V}{\partial t} = T_\alpha V_t$ on $\mathbb{R} \times \mathbb{R}^d$. Similarly, we have $\frac{\partial T_\alpha V}{\partial x_i} = T_\alpha V_{x_i}$ and $\frac{\partial T_\alpha V_{x_i}}{\partial x_j} = T_\alpha V_{x_i x_j}$ on $\mathbb{R} \times \mathbb{R}^d$ for $i, j = 1, \ldots, d$, which implies further that $\frac{\partial^2 T_\alpha V}{\partial x_i \partial x_j} = T_\alpha V_{x_i x_j}$ on $\mathbb{R} \times \mathbb{R}^d$.

Since the above sequence α' is arbitrary, for given $(\tilde{f}, \tilde{g}) \in H(f, g)$, we may assume that α' is such that $T_{\alpha'}(f, g) = (\tilde{f}, \tilde{g})$ uniformly on $\mathbb{R} \times S$ for any compact $S \subset \mathbb{R}^d$. So, in this case we have

$$T_\alpha(\mathscr{L}V) = \mathscr{L}_{T_\alpha(f,g)}(T_\alpha V) \quad \text{on } \mathbb{R} \times \mathbb{R}^d$$

with

$$\mathscr{L}_{T_\alpha(f,g)}(T_\alpha V) = \frac{\partial T_\alpha V}{\partial t}(t, x - y) + \sum_{i=1}^{d} (T_\alpha f_i(t,x) - T_\alpha f_i(t,y)) \frac{\partial T_\alpha V}{\partial x_i}(t, x - y)$$

$$+ \frac{1}{2} \sum_{i,j=1}^{d} \sum_{l=1}^{m} (T_\alpha g_{il}(t,x) - T_\alpha g_{il}(t,y)) \frac{\partial^2 T_\alpha V}{\partial x_i \partial x_j}(t, x - y)(T_\alpha g_{jl}(t,x) - T_\alpha g_{jl}(t,y)).$$

It is immediate to see that $T_\alpha V$ satisfies (h_1)-(h_2), and $\mathscr{L}_{T_\alpha(f,g)}(T_\alpha V) \geq a(|x - y|)$, i.e. (h_0) also holds with $T_\alpha V$, $T_\alpha f$, $T_\alpha g$ replacing V, f, g, respectively.

Step 3. Conclusion. Consider the hull equation

$$\mathrm{d}X = T_\alpha f(t, X)\mathrm{d}t + T_\alpha g(t, X)\mathrm{d}W.$$

Clearly the hypothesis **(H)** holds for this equation, so the same argument as in Step 1 implies that this hull equation admits a unique strong L^2-bounded solution on $\mathbb{R}$ for given Brownian motion W.

By the proof of Theorem 3.1, we know that if the unique strong L^2-bounded solution X of (1.3) satisfies $\|X\|_\infty \leq r_0$ for some $r_0 > 0$, then the unique L^2-bounded solutions of the hull equations are also bounded with the same r_0. Note that the pathwise uniqueness implies the uniqueness of laws on $\mathbb{R}^d$, so $|\mathcal{D}_{r_0}^{T_\alpha(f,g)}| = |\mathcal{D}_{r_0}^{(f,g)}| \leq 1$ and hence $|\cup_{r>0} \mathcal{D}_r^{T_\alpha(f,g)}| = |\cup_{r>0} \mathcal{D}_r^{(f,g)}| \leq 1$. The result now follows from Corollary 4.4. $\square$

Now we give some applications of Theorem 5.1.

Corollary 5.2. *Consider one dimensional linear stochastic differential equation with m-dimensional Brownian motion:*

$$(5.3) \qquad \mathrm{d}X(t) = (A(t)X(t) + f(t))\mathrm{d}t + \sum_{i=1}^{m} (B_i(t)X(t) + g_i(t))\mathrm{d}W_i(t),$$

where A, f, B_i, g_i are almost periodic functions. If there exists some constant $c > 0$ such that $2A(t) + \sum_{i=1}^{m} B_i^2(t) \geq c$ for all $t \in \mathbb{R}$, then all the L^2-bounded solutions of (5.3) share the same distribution on $\mathbb{R}$ which is almost periodic.

Proof. Consider the nonnegative function $V(t,x) = x^2 \exp\{\arctan t\}$ for $t \in \mathbb{R}$ and $x \in \mathbb{R}$. Note that (h_1)-(h_2) hold and the derivatives of V satisfy the boundedness condition on $\mathbb{R} \times S$ for any compact subset $S \subset \mathbb{R}$, which are required in Theorem 5.1. Since

$$\mathscr{L}V(t, x - y) = \frac{1}{t^2 + 1}V(t, x - y) + (2A(t) + \sum_{i=1}^{m} B_i^2(t))V(t, x - y)$$

$$\geq cV(t, x - y) \geq c\mathrm{e}^{-\pi/2}(x - y)^2,$$

the condition (h_0) holds. The result now follows from Theorem 5.1. $\square$

Corollary 5.3. *Consider one dimensional stochastic differential equation with m-dimensional Brownian motion:*

$$(5.4) \qquad \mathrm{d}X(t) = f(t, X(t))\mathrm{d}t + \sum_{i=1}^{m} g_i(t)\mathrm{d}W_i(t),$$

where g_i are almost periodic, $f(t,x)$ is uniformly almost periodic and global Lipschitz in x. If there is a constant $L_0 > 0$ such that

$$(5.5) \qquad L_0(x - y)^2 \leq (f(t,x) - f(t,y))(x - y) \quad \text{for all } t, x, y \in \mathbb{R},$$

then all the L^2-bounded solutions of (5.4) share the same distribution on $\mathbb{R}$ which is almost periodic.

Proof. Consider the function $V : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ given by

$$V(t, x) = \exp\{\arctan t\} \ln(x^2 + 1).$$

It is immediate to check that (h_1)-(h_2) hold and the derivatives of V satisfy the boundedness condition on $\mathbb{R} \times S$ for any compact subset $S \subset \mathbb{R}$. On the other hand,

$$\begin{aligned}
\mathscr{L}V(t, x - y) &= \frac{1}{t^2 + 1} V(t, x - y) + \frac{2(x - y)(f(t, x) - f(t, y)) \exp\{\arctan t\}}{(x - y)^2 + 1} \\
&\geq \frac{2L_0(x - y)^2 \exp\{\arctan t\}}{(x - y)^2 + 1}.
\end{aligned}$$

That is, (h_0) also holds. So the result follows from Theorem 5.1. $\qquad\square$

6. Some discussions and improvements

From the viewpoint of differential equations and dynamical systems, it is very natural to consider how the distribution of solutions evolves with the time, as we did in previous sections. Almost periodicity is an important recurrent property (in the sense of Birkhoff [4]), which roughly means that the motion will turn back repeatedly with any preassigned small error. When an equation, with recurrent solutions (motions), is stochastically perturbed, does the perturbed equation still admit recurrent motions in some sense? It is one of our main motivations to partly answer this problem; and it seems that it is appropriate, by the results in previous sections, to consider the recurrent motions in distribution sense on $\mathbb{R}^d$. However, some probabilists may prefer to consider properties they are interested in on path spaces, i.e. they think that properties for sample functions are more probabilistic. Due to this, in this section we discuss the almost periodicity of solutions on the path space; we note that the similar concept was considered by Da Prato and Tudor [12].

In Sections 3 and 4, we proved that under the Favard or Amerio separation condition (besides other conditions), equations (1.2) and (1.3) admit solutions which are almost periodic in distribution on $\mathbb{R}^d$. Indeed, under the Favard condition or the trivial Amerio separation condition, these solutions are almost periodic in distributions on the path space $C(\mathbb{R}, \mathbb{R}^d)$. It is well-known that $C(\mathbb{R}, \mathbb{R}^d)$ is a separable complete metric space with the metric

$$d(\omega_1, \omega_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min\left\{1, \sup_{-n \leq t \leq n} |\omega_1(t) - \omega_2(t)|\right\},$$

i.e. the convergence on the path space means the uniform convergence on any compact interval.

For any solution X of (1.3), it determines a distribution on $C(\mathbb{R}, \mathbb{R}^d)$. Denote the shift mapping

$$\hat{\mu} : \mathbb{R} \to \mathcal{P}(C(\mathbb{R}, \mathbb{R}^d)), \ t \mapsto \hat{\mu}(t) := \mathcal{L}(X(t + \cdot)),$$

where $\mathcal{P}(C(\mathbb{R}, \mathbb{R}^d))$ stands for the space of probability measures on the path space and $\mathcal{L}(X(t + \cdot))$ means the law of the $C(\mathbb{R}, \mathbb{R}^d)$-valued random variable $X(t + \cdot)$. Note that $\mathcal{P}(C(\mathbb{R}, \mathbb{R}^d))$ is a separable complete metric space (see, e.g. [32, Chapter II, Theorems 6.2 and 6.5]). The solution X is said to be *almost periodic in strong distribution sense* if $\hat{\mu}$ is a $\mathcal{P}(C(\mathbb{R}, \mathbb{R}^d))$-valued almost periodic mapping. It is clear that if X is almost periodic in strong distribution sense, then it is almost periodic in distribution.

Firstly the result of Theorem 3.1 can be improved.

Theorem 3.1'. *Consider the following family of Itô stochastic equations on $\mathbb{R}^d$*

$$dX = f_n(t, X)dt + g_n(t, X)dW, \quad n = 1, 2, \cdots,$$

where f_n are $\mathbb{R}^d$-valued, g_n are $(d \times m)$-matrix-valued, and W is a standard m-dimensional Brownian motion. Assume that f_n, g_n satisfy the conditions of global Lipschitz and linear

growth with common Lipschitz and linear growth constants; that is, there are constants L and K, independent of $t \in \mathbb{R}$ and $n \in \mathbb{N}$, such that for all $x, y \in \mathbb{R}^d$

$$|f_n(t,x) - f_n(t,y)| \vee |g_n(t,x) - g_n(t,y)| \le L|x - y|,$$
$$|f_n(t,x)| \vee |g_n(t,x)| \le K(1 + |x|).$$

Assume further that $f_n \to f$, $g_n \to g$ pointwise on $\mathbb{R} \times \mathbb{R}^d$ as $n \to \infty$ and that $X_n \in \mathcal{B}_{r_0}^{(f_n, g_n)}$ for some constant r_0, independent of n. Then there is a subsequence of $\{X_n\}$ which converges, in strong distribution sense (the meaning is obvious), to some $X \in \mathcal{B}_{r_0}^{(f,g)}$.

Proof. We only need to point out the difference from the proof of Theorem 3.1. Note that, in the proof of Theorem 3.1, (3.2) actually implies that $\tilde{X}_n$ converges in distribution to $\tilde{X}$ on $C([a,b], \mathbb{R}^d)$. Since the interval $[a,b]$ is arbitrary, the convergence indeed occurs on $C(\mathbb{R}, \mathbb{R}^d)$ in distribution. $\qquad \square$

By Theorem 3.1' and minor revising the proof of Lemma 3.10, we have the following result.

Lemma 3.10'. *Assume that each hull equation of (1.3) admits a unique minimal solution in strong distribution sense, i.e. all the minimal solutions of the given hull equation possess the same law on the path space. Then these minimal solutions are almost periodic in strong distribution sense.*

Therefore, we have the following

Theorem A'. *Consider (1.2) with the coefficients $A, B_1, \ldots, B_m$, $f, g_1, \ldots, g_m$ being almost periodic. Assume further that (1.2) admits an L^2-bounded solution X, and that the Favard condition holds for (1.2). Then (1.2) admits a solution which is almost periodic in strong distribution sense.*

Proof. The theorem follows immediately from Lemmas 3.3, 3.6, Remark 3.7, Corollary 3.9, and Lemma 3.10'. $\qquad \square$

Remark 6.1. Note that Remark 3.12 and Corollary 3.13 can be correspondingly improved by Theorem A'.

For the nonlinear equation (1.3), under the trivial Amerio separation condition, we have the following

Corollary 4.4'. *Consider (1.3). Assume (**H**) and that each hull equation $(\tilde{f}, \tilde{g})$ of (1.3) admits, in $\mathcal{B}_r^{(\tilde{f}, \tilde{g})}$, a unique distribution in strong sense, i.e. all the elements of $\mathcal{B}_r^{(\tilde{f}, \tilde{g})}$ share the same distribution on the path space. Then $\mathcal{B}_r^{(1.3)}$ consists of solutions which are almost periodic in strong distribution sense with the unique common distribution on the path space.*

Proof. The proof is completely similar to that of Lemma 3.10' since only the uniqueness of distribution on the path space is essential in the proof. $\qquad \square$

Remark 6.2. We note that the conclusion of Theorem 5.1 (and hence its applications) can be correspondingly improved by Corollary 4.4'.

Acknowledgements

References

[1] L. Amerio, Soluzioni quasi-periodiche, o limitate, di sistemi differenzali non lineari quasi-periodiche, o limitati, *Ann. Mat. Pura Appl.* **39** (1955), 97–119.

[2] L. Amerio and G. Prouse, Almost-periodic Functions and Functional Equations. Van Nostrand Reinhold Co., New York-Toronto, Ont.-Melbourne 1971 viii+184 pp.

[3] L. Arnold and C. Tudor, Stationary and almost periodic solutions of almost periodic affine stochastic differential equations, *Stochastics Stochastics Rep.* **64** (1998), 177–193.

[4] G. D. Birkhoff, Dynamical Systems. American Mathematical Society Colloquium Publications, Vol. IX, American Mathematical Society, Providence, R.I. 1927.

[5] S. Bochner, Beiträge zur theorie der fastperiodischen funktionen, I. Funktionen einer Variablen (German), *Math. Ann.* **96** (1927), 119–147.

[6] S. Bochner, Abstrakte fastperiodische funktionen (German), *Acta Math.* **61** (1933), 149–184.

[7] S. Bochner, A new approach to almost periodicity, *Proc. Nat. Acad. Sci. U.S.A.* **48** (1962), 2039–2043.

[8] H. Bohr, Zur theorie der fastperiodischen funktionen. (German) I. *Acta Math.* **45** (1924), 29–127.

[9] H. Bohr, Zur theorie der fastperiodischen funktionen. (German) II. *Acta Math.* **46** (1925), 101–214.

[10] H. Bohr, Zur theorie der fastperiodischen funktionen. (German) III. *Acta Math.* **47** (1926), 237–281.

[11] H. Bohr and O. Neugebauer, Über lineare Differentialgleichungen mit konstanten Koeffizienten und fast-periodischer rechter Seite, *Nachr. Ges. Wiss. Göttingen, Math.-Phys.* Klasse (1926), 8–22.

[12] G. Da Prato and C. Tudor, Periodic and almost periodic solutions for semilinear stochastic equations, *Stochastic Anal. Appl.* **13** (1995), 13–33.

[13] R. M. Dudley, Real Analysis and Probability. Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics, 74. Cambridge University Press, Cambridge, 2002. x+555 pp.

[14] J. Favard, Sur les équations diffrentielles linéaires à coefficients presque-périodiques (French), *Acta Math.* **51** (1927), 31–81.

[15] J. Favard, Lecons sur les Fonctions Presque-Périodiques, Gauthier-Villars, Paris, 1933.

[16] A. M. Fink, Semi-separated conditions for almost periodic solutions, *J. Differential Equations* **11** (1972), 245–251.

[17] A. M. Fink, Almost Periodic Differential Equations. Lecture Notes in Mathematics, Vol. 377. Springer-Verlag, Berlin-New York, 1974. viii+336 pp.

[18] A. Friedman, Stochastic Differential Equations and Applications. Vol. 1. Probability and Mathematical Statistics, Vol. 28. Academic Press, New York-London, 1975. xiii+231 pp.

[19] I. I. Gihman and A. V. Skorohod, Stochastic Differential Equations. Translated from the Russian by Kenneth Wickwire. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72. Springer-Verlag, New York-Heidelberg, 1972. viii+354 pp.

[20] A. Halanay, Periodic and almost periodic solutions to affine stochastic systems. Proceedings of the Eleventh International Conference on Nonlinear Oscillations (Budapest, 1987), 94–101, János Bolyai Math. Soc., Budapest, 1987.

[21] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes. Second edition. North-Holland Mathematical Library, 24. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. xvi+555 pp.

[22] R. A. Johnson, A linear almost periodic equation with an almost automorphic solution, *Proc. Amer. Math. Soc.* **82** (1981), 199–205.

[23] E. R. van Kampen, Almost periodic functions and compact groups, *Ann. of Math. (2)* **37** (1936), 78–91.

[24] I. Karatzas and S. E. Shreve, Brownian Motion and Stochastic Calculus. Second edition. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991. xxiv+470 pp.

[25] J. L. Kelley, General Topology. Reprint of the 1955 edition [Van Nostrand, Toronto, Ont.]. Graduate Texts in Mathematics, No. 27. Springer-Verlag, New York-Berlin, 1975. xiv+298 pp.

[26] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms. École d'été de probabilités de Saint-Flour, XII1982, 143–303, Lecture Notes in Math., 1097, Springer, Berlin, 1984.

[27] B. M. Levitan and V. V. Zhikov, Almost Periodic Functions and Differential Equations. Translated from the Russian by L. W. Longdon. Cambridge University Press, Cambridge-New York, 1982. xi+211 pp.

[28] J. L. Massera, The existence of periodic solutions of systems of differential equations, *Duke Math. J.* **17** (1950), 457–475.

[29] T. Morozan and C. Tudor, Almost periodic solutions of affine Itô equations, *Stoch. Anal. Appl.* **7** (1989), 451–474.

[30] J. von Neumann, Almost periodic functions in a group. I. *Trans. Amer. Math. Soc.* **36** (1934), 445–492.

[31] R. Ortega and M. Tarallo, Almost periodic linear differential equations with non-separated solutions, *J. Funct. Anal.* **237** (2006), 402–426.

[32] K. R. Parthasarathy, Probability Measures on Metric Spaces. Probability and Mathematical Statistics, No. 3 Academic Press, Inc., New York-London 1967 xi+276 pp.

[33] Y. V. Prohorov, Convergence of random processes and limit theorems in probability theory, *Theory Probab. Appl.* **1** (1956), 157–214.

[34] G. Seifert, A condition for almost periodicity with some applications to functional-differential equations, *J. Differential Equations* **11** (1965), 393–408.

[35] G. Seifert, On uniformly almost periodic sets of functions for almost periodic differential equations, *Tohoku Math. J.* **34** (1982), 301–309.

[36] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations, *J. Math. Kyoto Univ.* **11** (1971), 155–167.

[37] T. Yoshizawa, Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions. Applied Mathematical Sciences, Vol. 14. Springer-Verlag, New York-Heidelberg, 1975. vii+233 pp.

[38] V. V. Zhikov and B. M. Levitan, Favard theory, *Russian Math. Surveys* **32** (1977), 129–180.

Z. Liu: School of Mathematics, Jilin University, Changchun 130012, P. R. China and School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, P. R. China
E-mail address: `zxliu@jlu.edu.cn; zxliu@dlut.edu.cn`

W. Wang: School of Mathematics, Jilin University, Changchun 130012, P. R. China
E-mail address: `whwang13@mails.jlu.edu.cn`